CLAIMS

1.   A method comprising the steps of:

receiving data from a web site located across a network;

determining whether additional data from the web site is extractable;

in response to determining that additional data from the web site is extractable, creating at least one synthetic hyperlink for extracting the data from the web site;

combining the at least one synthetic hyperlink with the data received from the website to create combined data; and

sending the combined data to a crawler.

2.   The method of claim 1, wherein the step of creating at least one synthetic hyperlink comprises the step of:

augmenting a hyperlink request with at least one parameter value appropriate for a form used by the web site to create at least one synthetic hyperlink.

3.   The method of claim 1, wherein the step of creating at least one synthetic hyperlink comprises the step of:

augmenting a hyperlink request with at least one parameter value determined by executing a script contained in the data from the web site to create at least one synthetic hyperlink.

4.   The method of claim 1, wherein the step of creating at least one synthetic hyperlink comprises the step of:

augmenting a hyperlink request with at least one parameter value determined by a script filter analyzing a script contained in the data from the web site to create at least one synthetic hyperlink.

5.   The method of claim 1, wherein the step of creating at least one synthetic hyperlink comprises the step of:

augmenting a hyperlink request with at least one parameter value to create a synthetic hyperlink, the synthetic hyperlink indicating that it must be converted to a different method for extracting data from the web site.

6.   A method comprising the steps of:

receiving a synthetic hyperlink request;

converting the synthetic hyperlink request to a method indicated by the synthetic hyperlink request to create a converted hyperlink request; and

5   sending the converted hyperlink request to a web site.

7.   The method of claim 6, wherein the step of converting the synthetic hyperlink request comprises the steps of:

determining whether the synthetic hyperlink request indicates the use of a

10   POST method; and

in response to determining that the synthetic hyperlink request indicates the use of a POST method, converting the synthetic hyperlink from a GET method to a POST method.

15

8.   A system comprising:

at least one communication device for receiving data from a web site and sending data to a crawler; and

at least one computer processing device, communicatively coupled to the

5   at least one communication device, for

analyzing data received from the web site to determine whether additional data is extractable from the web site by a hyperlink request,

creating at least one synthetic hyperlink for extracting data from the web site, and

10   combining the at least one synthetic hyperlink with data received from the web site in response to the hyperlink request.

9.   The system according to claim 8, wherein the at least one computer processing device comprises:

15   a transcoder for

determining at least one parameter value for a form required by the data from the web site;

augmenting the hyperlink request with at least one parameter value to create the at least one synthetic hyperlink; and

20   combining the at least one synthetic hyperlink with the data received from the web site in response to the hyperlink request.

10.   The system according to claim 8, wherein the at least one computer processing device comprises:

a transcoder, communicatively coupled to the at least one communication device, for

augmenting the hyperlink request with at least one

parameter value to create the at least one synthetic hyperlink for

each of the at least one parameter value, and

combining the at least one synthetic hyperlink with the data

received from the web site; and

a script engine, communicatively coupled to the transcoder, for determining the at least one parameter value by executing a script contained in the data from the web site.

11.   The system according to claim 8, wherein the at least one computer processing device comprises:

a transcoder, communicatively coupled to the at least one communication device, for

augmenting the hyperlink request with at least one

parameter value to create the at least one synthetic hyperlink for

each of the at least one parameter value, and

combining the at least one synthetic hyperlink with the data

received from the web site; and

a script filter, communicatively coupled to the transcoder, for determining the at least one parameter value by analyzing the script contained in the data from the web site.

12.   The system according to claim 8, wherein the at least one computer processing device comprises:

a transcoder for augmenting the hyperlink request with at least one parameter value to create the at least one synthetic hyperlink indicating that the at least one synthetic hyperlink must be converted to a different method for extracting data from the web site.

13.   A system for extending the performance of a crawler comprising:

at least one communication device for receiving a synthetic hyperlink request and sending data to a web site across a network, in accordance with the synthetic hyperlink request; and

a method converter, communicatively coupled to the at least one communication device, for converting the synthetic hyperlink request to a method compatible with the web site.

14.   The system of claim 13, wherein the method converter is for converting the synthetic hyperlink request from a GET method to a POST method.

15.   A computer readable medium including computer instructions for a computing system, the computer instructions comprising instructions for:

receiving data from a web site located across a network;

determining whether additional data from the web site is extractable;

in response to determining that additional data from the web site is extractable, creating at least one synthetic hyperlink for extracting the data from the web site;

combining the at least one synthetic hyperlink with the data received from the web site to create combined data; and

sending the combined data to a crawler.

16.   The computer readable medium of claim 15, wherein the computer instructions for creating at least one synthetic hyperlink comprise instructions for augmenting a hyperlink request with at least one parameter value for a form used by the web site to create the at least one synthetic hyperlink.

17.   The computer readable medium of claim 15, wherein the computer instructions for creating at least one synthetic hyperlink comprise instructions for augmenting a hyperlink request with at least one parameter value determined by executing a script contained in the data from the web site to create the at least one synthetic hyperlink.

18.   The computer readable medium of claim 15, wherein the computer instructions for creating at least one synthetic hyperlink comprise instructions for augmenting a hyperlink request with at least one parameter value determined by a script filter by analyzing a script contained in the data from the web site to create the at least one synthetic hyperlink.

Docket No. ARC920000141US1   - 26 -

19.   A computer readable medium including computer instruction for a
computer system, the computer instructions comprising instructions for:
    receiving a synthetic hyperlink request from a crawler;
    converting the synthetic hyperlink request to a method indicated by the
synthetic hyperlink request to create a converted hyperlink request; and
    sending the converted hyperlink request to a web site.

20.   A computer readable medium including computer instruction for a
computer system, the computer instructions comprising instructions for:

      receiving a synthetic hyperlink request for extracting data from a web site;

      converting the synthetic hyperlink request to a method indicated by the
synthetic hyperlink request; and

      in response to determining that the synthetic hyperlink request indicates
the use of a POST method, converting the synthetic hyperlink request from a
GET method to a POST method.